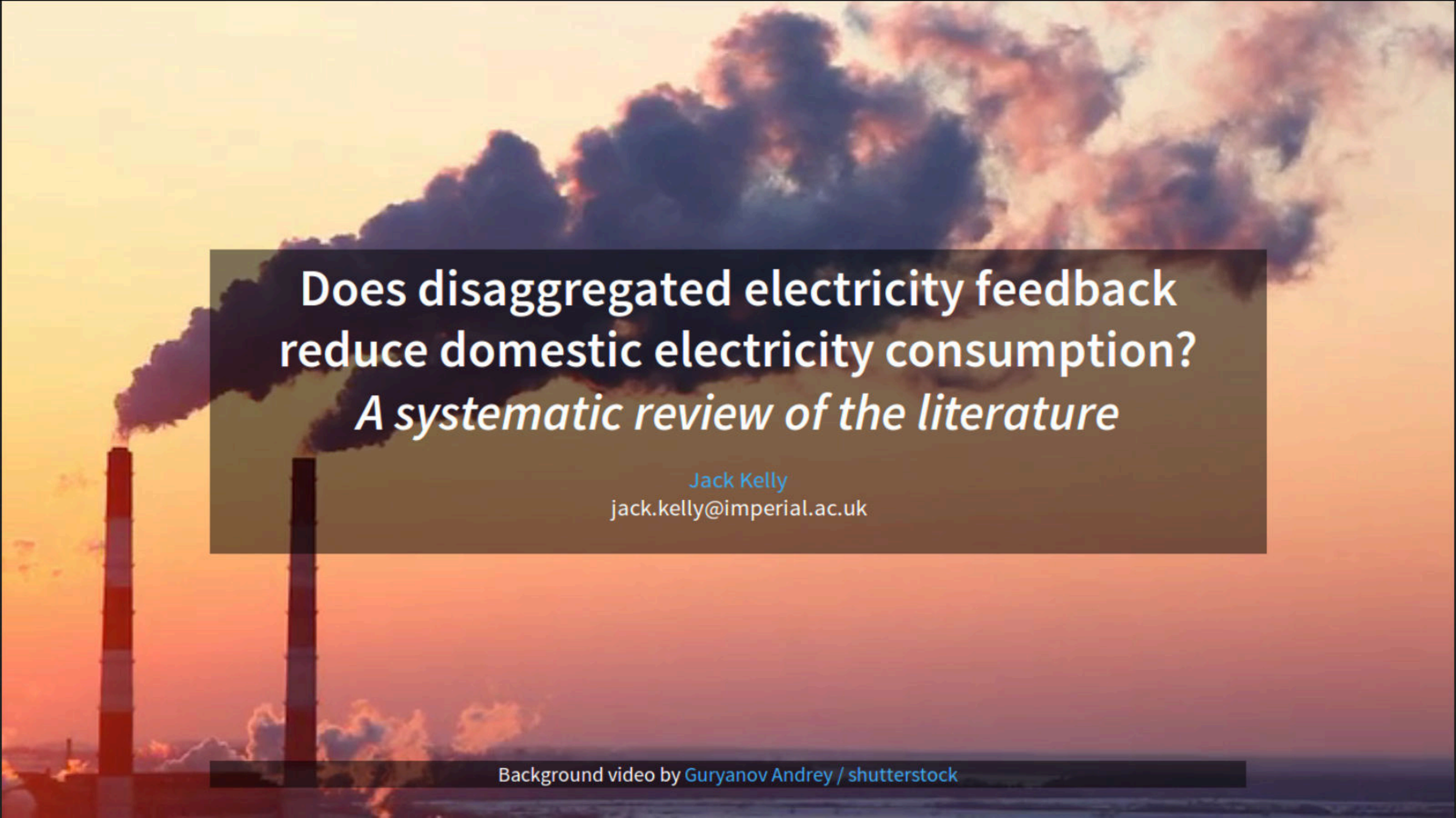


Does disaggregated electricity feedback reduce domestic electricity consumption?
A systematic review of the literature

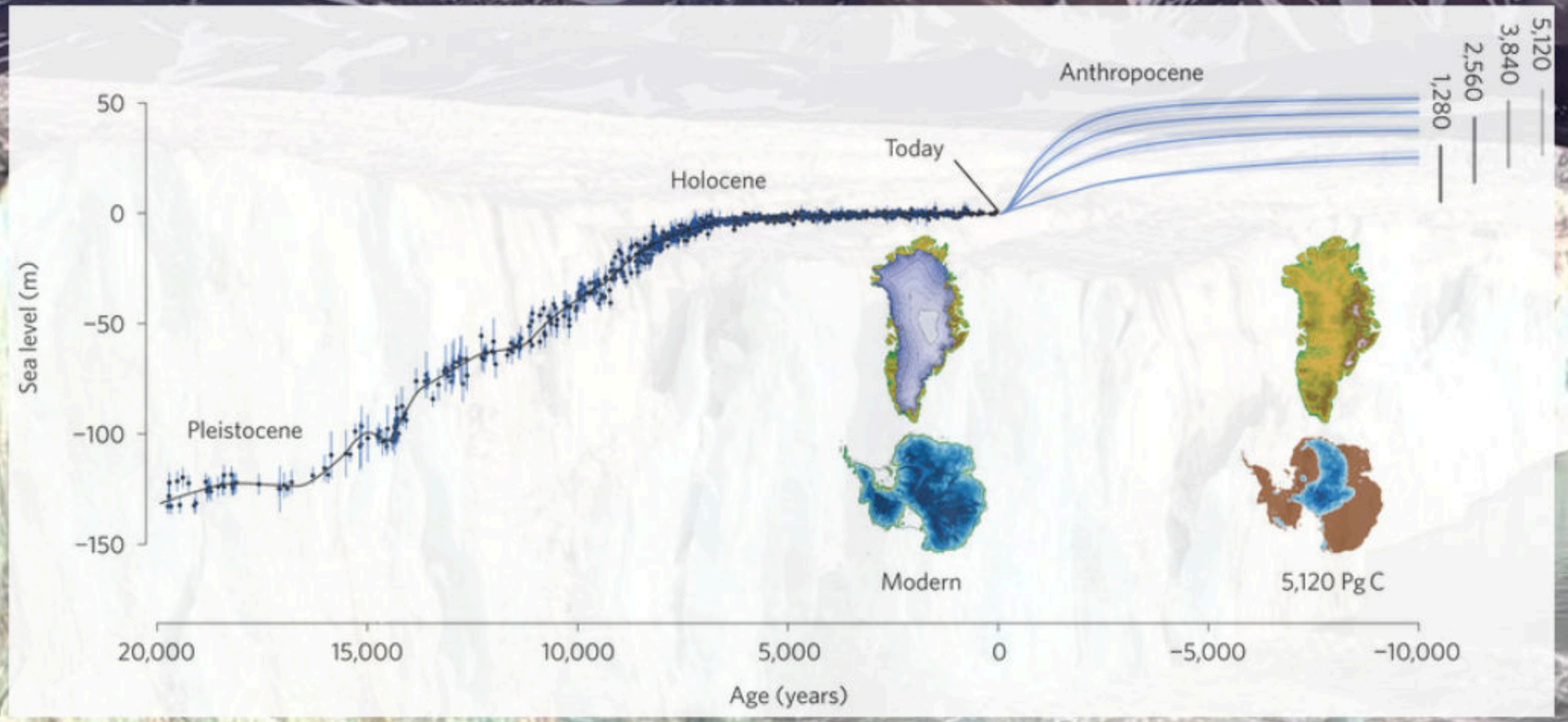
Jack Kelly
jack.kelly@imperial.ac.uk

Background video by Guryanov Andrey / shutterstock


Background video by Incredible Arctic / shutterstock


Background video by Incredible Arctic / shutterstock

Past and future changes in global mean sea level



Clark et al., Consequences of twenty-first-century policy for multi-millennial climate and sea-level change, *Nature Climate Change*, 2016

Background video by Incredible Arctic / shutterstock

Evidence that NILM can help save energy...

Evidence that NILM can help save energy...

1) People *want* disaggregated energy data

2) Behaviour affects energy consumption

2) Behaviour affects energy consumption

modifying behaviour → reduce energy consumption

3) People are bad at estimating
the energy consumption of their appliances

3) People are bad at estimating the energy consumption of their appliances

→ Fix the ‘information deficit’ then users can operate as rational ‘resource managers’

3) People are bad at estimating the energy consumption of their appliances

→ Fix the ‘information deficit’ then users can operate as rational ‘resource managers’

(*I’m now sceptical of this idea*)

4) Multiple studies *report* that disaggregated feedback reduces energy consumption

5) Smart meters

‘NILM

dead!’

‘NILM is dead!’

Systematic reviews

Background image from UCSF

Systematic reviews

- Common in medicine, social sciences etc.
- Distinct from ‘narrative’ reviews
- Aim to collect *all* papers matching a defined search criteria
- Quantitative summary of each paper and biases
- Quantitative synthesis of all results

Background image from UCSF

Literature search

1. Three search engines: Google Scholar, the ACM Digital Library and IEEE Xplore
2. Search terms:
   - ‘disaggregated AND [energy|electricity] AND feedback’
   - ‘N[I|A|IA]LM AND feedback’
3. Searched papers’ bibliographies
4. Sent draft literature review to authors for comments

The studies

12 groups of studies identified

Study	Feedback presentation	Num. houses in disag. group	Num. houses in study	Num. disaggregation categories	Duration (months of disag)	Reduction in electricity use[U] (%)	Reduction is for whole house?	Sample period of meter	Feedback delay	Timing: Historic or Concurrent?	Time frames for historic[T]	Recommendations given?[R]	Control group?	Controlled for Hawthorne?	Volunteer bias?[V]	Controlled for weather?
"RECS" [23]	dedicated computer	25	100	~ 8	2	12.9	✓	0.6 sec	0	H&C	HDM	✗	✓	✓	L	✓
McCalley & Midden 2002 [24]	Virt. wash. machine	25	100	1	-	0.0	✗	-	0	H&C	-	G	✓	✓	L	-
Wood & Newborough '03 [25]; Mansouri & Newborough '99 [26]	LCD by cooker	10	44	1	≥ 2	12.2	✗	15 sec	0	C	-	✗	✓	✓	L	✓
"ECOIS-I" [27], [28]	Dedicated laptop	8	8	16	2	9	✓	30 min	next day	H	D, 10D	P	✗	✗	H[#]	✓
"ECOIS-II" [28]–[30]	Dedicated laptop	10	19	16	3	18	✓	30 min	next day	H	D, 10D	P	✓	✓	H[#]	✓
"EnergyLife" trial 1 [31]–[33]	iPhone	13	13	7	3	5	✓	?	1-2 min	H&C	D	P	✗[#]	✗[#]	H[#]	✗[#]
"EnergyLife" trial 2 [34]	iPhone	4	4	7	4	38	✗	?	1-2 min	H&C	D	P	✗	✗[#]	H[#]	✗[#]
Home Energy Analytics [15], [16], [20], [21]	Web & email & home visits	1623	1623	5	≤ 44	6.1	✓	hourly	0	H	Y	P	✗	✗	L	✓
Bidgely 2013 [35], [36]	Web, mobile, email	163	328	≥ 3?	-	6	✓	30 sec & 1 hr	0	H&C	DBY	P	✓	✗	H	✓
PG&E Pilot 2014 [17], [22]	Web, mobile, email	844	1685	≥ 3?	3	2.1	✓	30 sec	0	H&C	DBY	P	✓	✗	H	✓
Schwartz et al. 2014 [14]	Web, mob, TV	6	6	~ 10	18	7.8	✓	?	0?	H&C	?	?	✗	✗	H	✗
Sokoloski 2015 [37]	Web, mob, email	12	70	≥ 3?	0.75	3	✓	30 sec	0	H&C	DBY	P	✓	✗	L	✓

Research questions

Q1. Can disaggregated electricity feedback enable ‘energy enthusiasts’ to save energy?

Q1. Can disaggregated electricity feedback enable ‘energy enthusiasts’ to save energy?

- Very likely...
- Weighted-mean energy reduction = 4.5%
- A *lot* of uncertainty...

Biases

The Hawthorne Effect

The Hawthorne Effect

- Hawthorne effect is illustrated by Schwartz et al. 2013:

The Hawthorne Effect

- Hawthorne effect is illustrated by Schwartz et al. 2013:
  - Randomised controlled trial

The Hawthorne Effect

- Hawthorne effect is illustrated by Schwartz et al. 2013:
  - Randomised controlled trial
  - 6,350 participants split into 2 groups: control & treatment

The Hawthorne Effect

- Hawthorne effect is illustrated by Schwartz et al. 2013:
  - Randomised controlled trial
  - 6,350 participants split into 2 groups: control & treatment
  - Treatment received weekly postcard saying: '*You have been selected to be part of a one-month study of how much electricity you use in your home... No action is needed on your part. We will send you a weekly reminder postcard about the study...*'

The Hawthorne Effect

- Hawthorne effect is illustrated by Schwartz et al. 2013:
  - Randomised controlled trial
  - 6,350 participants split into 2 groups: control & treatment
  - Treatment received weekly postcard saying: '*You have been selected to be part of a one-month study of how much electricity you use in your home... No action is needed on your part. We will send you a weekly reminder postcard about the study...*'
  - Treatment group reduced energy consumption by 2.7%!

The Hawthorne Effect

- Hawthorne effect is illustrated by Schwartz et al. 2013:
  - Randomised controlled trial
  - 6,350 participants split into 2 groups: control & treatment
  - Treatment received weekly postcard saying: '*You have been selected to be part of a one-month study of how much electricity you use in your home... No action is needed on your part. We will send you a weekly reminder postcard about the study...*'
  - Treatment group reduced energy consumption by 2.7%!
- Failure to control for Hawthorne very likely to be strong positive bias

The Hawthorne Effect

- Hawthorne effect is illustrated by Schwartz et al. 2013:
  - Randomised controlled trial
  - 6,350 participants split into 2 groups: control & treatment
  - Treatment received weekly postcard saying: '*You have been selected to be part of a one-month study of how much electricity you use in your home... No action is needed on your part. We will send you a weekly reminder postcard about the study...*'
  - Treatment group reduced energy consumption by 2.7%!
- Failure to control for Hawthorne very likely to be strong positive bias
- 8 studies did not control for Hawthorne

Other biases

Other biases

- 6 studies used attention-grabbing displays
- Home-visits
- 10 studies were short (4 months or less)
- Cherry-picking statistical analyses and comparison periods?
- 8 studies used sub-metered data, hence avoiding mistrust from participants
- Publication bias?

Q2. How much energy would the *whole* population save?

Q2. How much energy would the *whole* population save?

- All 12 studies suffer from ‘opt-in’ bias

Q2. How much energy would the *whole* population save?

- All 12 studies suffer from ‘opt-in’ bias
- Subjects self-selected hence are probably more interested in energy than the average person

Q2. How much energy would the *whole* population save?

- All 12 studies suffer from ‘opt-in’ bias
- Subjects self-selected hence are probably more interested in energy than the average person
- Very likely to be a strong positive bias

Q3. Aggregate versus disaggregated feedback

Q3. Aggregate versus disaggregated feedback

- 4 of the 12 studies directly compared disaggregated against aggregate feedback

Q3. Aggregate versus disaggregated feedback

- 4 of the 12 studies directly compared disaggregated against aggregate feedback
  - 3 studies found aggregate to be *more* effective

Q3. Aggregate versus disaggregated feedback

- 4 of the 12 studies directly compared disaggregated against aggregate feedback
  - 3 studies found aggregate to be *more* effective
  - 1 study found aggregate to be *equally* effective

Q3. Aggregate versus disaggregated feedback

- 4 of the 12 studies directly compared disaggregated against aggregate feedback
  - 3 studies found aggregate to be *more* effective
  - 1 study found aggregate to be *equally* effective
  - 2 field trials & 2 lab experiments

The 2 field trials...

Sokoloski 2015

Sokoloski 2015

- Randomised controlled trial

Sokoloski 2015

- Randomised controlled trial
- 70 households recruited in California

Sokoloski 2015

- Randomised controlled trial
- 70 households recruited in California
- 3 weeks

Sokoloski 2015

- Randomised controlled trial
- 70 households recruited in California
- 3 weeks
- Rebecca Sokoloski, Disaggregated Electricity Consumption: Using Appliance-Specific Feedback to Promote Energy Conservation, M.A. thesis in Psychology, *California State University San Marcos*, 2015


bidgely
24 HR
BILLING CYCLE
YEAR
THU SEP 4, 2014
$
You spent $0.80 on Whole House
$0.10
$0.05
$0.00
10 11 12P 1 2 3 4 5 6 7 8 9 10 11 Thu 1 2 3 4 5 6 7 8 9 10
TIER1
NOW
0.14kW
WHOLE HOUSE $0.80
REFRIGERATION $0.24
ALWAYS ON $0.17
OTHERS $0.39
Bidgely Detective BETA
How much are your appliances costing you each month?
Find out yourself
Whole House usage details
ENERGY RATE
$0.16/kWh
YOUR CURRENT RATE
TIERS
1
$0.16
COST PROJECTIONS
$26
ESTIMATED BILL
CURRENT VS. PROJECTION
$5
$26
25 days left
© Bidgely 2014
Terms
Privacy
Go To Old Dashboard

bidgely

24 HR | BILLING CYCLE | YEAR

JUL 29 - AUG 28 2014

$

You spent $3.01 on Refrigeration

WHOLE HOUSE	$27.03
ALWAYS ON	$5.27
REFRIGERATION	$3.01
OTHERS	$18.76

Bidgely Detective *BETA*

How much are your appliances costing you each month?

Find out yourself

Refrigeration usage details

11%

OF WHOLE HOUSE

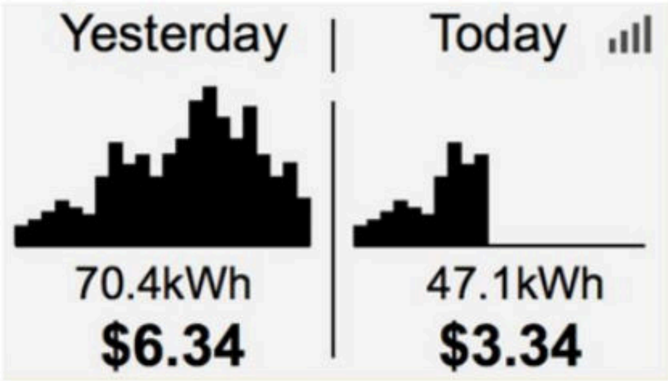
REFRIGERATION: $3.01 WHOLE HOUSE: $27.03

YOUR REFRIGERATION COMPARED TO SIMILAR

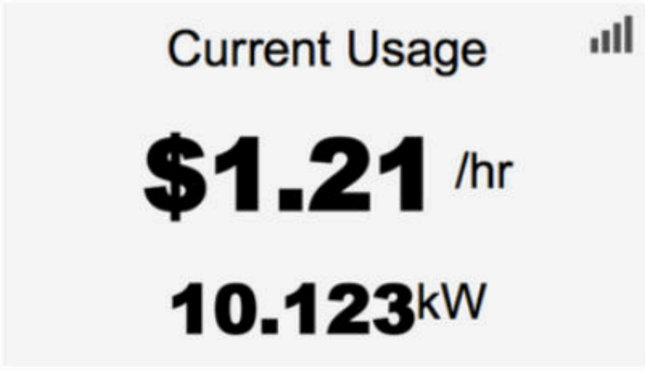
 Terms Privacy

Go To Old Dashboard

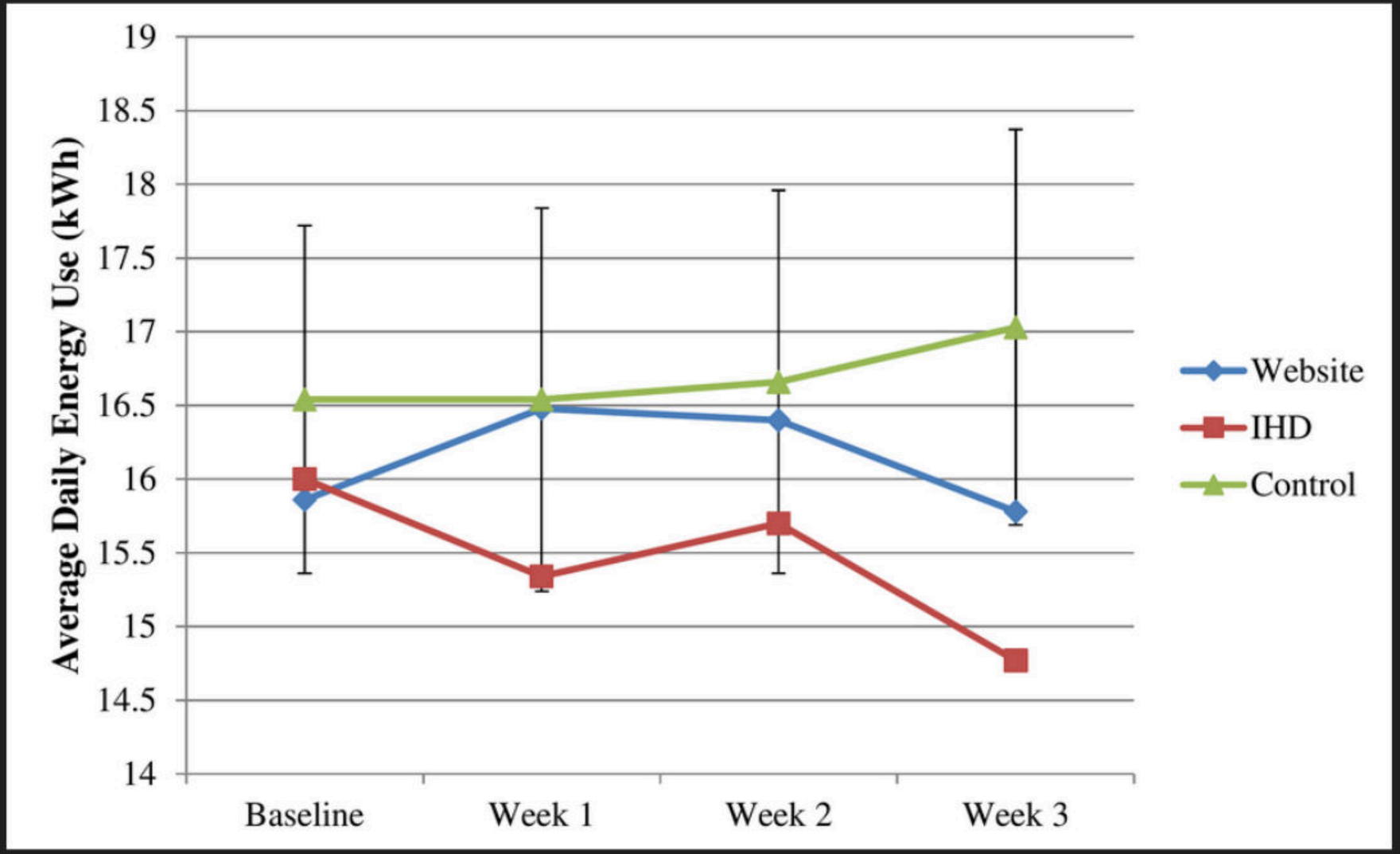
Appendix E.1: Day by Day Comparisons



Appendix E.2: Energy use in real-time



Sokoloski's results



Sokoloski's results

Energy reductions:

- IHD: 8.1% (statistically significant)
- Disaggregation: 0.5%
- Control: -2.5%

Sokoloski’s results

Findings from surveys:

Sokoloski's results

Findings from surveys:

- Follow-up survey revealed that the disag group were *not* significantly more likely to be willing to replace large, inefficient appliances compared to controls or IHD group.

Sokoloski's results

Findings from surveys:

- Follow-up survey revealed that the disag group were *not* significantly more likely to be willing to replace large, inefficient appliances compared to controls or IHD group.
- Neither controls nor the disag group significantly increased their perception of control (initial survey versus follow-up).

Sokoloski's results

Findings from surveys:

- Follow-up survey revealed that the disag group were *not* significantly more likely to be willing to replace large, inefficient appliances compared to controls or IHD group.
- Neither controls nor the disag group significantly increased their perception of control (initial survey versus follow-up).
- IHD group *did* increase their perception of control.

Sokoloski’s results

Findings from surveys:

Sokoloski's results

Findings from surveys:

- Users viewed their devices:
  - 0.86 times per day for disag users
  - 8.16 times per day for IHD users

PG&E 2014 trial

PG&E 2014 trial

- 1,685 PG&E customers

PG&E 2014 trial

- 1,685 PG&E customers
- 3 months

PG&E 2014 trial

- 1,685 PG&E customers
- 3 months
- Half got IHD & half got Bidgely

PG&E 2014 trial

- 1,685 PG&E customers
- 3 months
- Half got IHD & half got Bidgely
- Users choose intervention

PG&E 2014 trial

- 1,685 PG&E customers
- 3 months
- Half got IHD & half got Bidgely
- Users choose intervention
- Did not tease apart consumption of IHD vs Bidgely

PG&E 2014 trial

- 1,685 PG&E customers
- 3 months
- Half got IHD & half got Bidgely
- Users choose intervention
- Did not tease apart consumption of IHD vs Bidgely
- Churchwell et al., HAN Phase 3 Impact and Process Evaluation Report, technical report by Nexant, 2014

PG&E 2014 trial results

PG&E 2014 trial results

- IHD users significantly more likely to report taking actions to reduce electricity usage and to use their device to deduce power demand of individual appliances(!)

PG&E 2014 trial results

- IHD users significantly more likely to report taking actions to reduce electricity usage and to use their device to deduce power demand of individual appliances(!)
- Several users did not trust the disag data.

PG&E 2014 trial results

- IHD users significantly more likely to report taking actions to reduce electricity usage and to use their device to deduce power demand of individual appliances(!)
- Several users did not trust the disag data.
- IHD more successful in communicating power demand *now*

Bidgely have redesigned their website since these studies

Conclusions

Conclusions

- NILM has *many* uses! This talk just considered one use!

Conclusions

- NILM has *many* uses! This talk just considered one use!
- Available evidence suggests that aggregate feedback is *more effective* than disag feedback

Conclusions

- NILM has *many* uses! This talk just considered one use!
- Available evidence suggests that aggregate feedback is *more effective* than disag feedback
- But these results confounded by effect of IHD versus website

Conclusions

- NILM has *many* uses! This talk just considered one use!
- Available evidence suggests that aggregate feedback is *more effective* than disag feedback
- But these results confounded by effect of IHD versus website
- Disag feedback might drive savings of 0.7% - 4.5% in general population

Conclusions

- NILM has *many* uses! This talk just considered one use!
- Available evidence suggests that aggregate feedback is *more effective* than disag feedback
- But these results confounded by effect of IHD versus website
- Disag feedback might drive savings of 0.7% - 4.5% in general population
- Disag feedback might drive larger savings in ‘energy enthusiast’ populations

Conclusions

- NILM has *many* uses! This talk just considered one use!
- Available evidence suggests that aggregate feedback is *more effective* than disag feedback
- But these results confounded by effect of IHD versus website
- Disag feedback might drive savings of 0.7% - 4.5% in general population
- Disag feedback might drive larger savings in ‘energy enthusiast’ populations
- Fine-grained disag may not be necessary

Conclusions

- NILM has *many* uses! This talk just considered one use!
- Available evidence suggests that aggregate feedback is *more effective* than disag feedback
- But these results confounded by effect of IHD versus website
- Disag feedback might drive savings of 0.7% - 4.5% in general population
- Disag feedback might drive larger savings in ‘energy enthusiast’ populations
- Fine-grained disag may not be necessary
- But! Lots of gaps in our knowledge. Cannot robustly falsify any hypotheses yet.

Suggestions for future studies

Suggestions for future studies

- Compare aggregate versus disagg (both on an IHD)

Suggestions for future studies

- Compare aggregate versus disagg (both on an IHD)
- Compare 2 groups:
  1. Aggregate on an IHD
  2. Aggregate (on an IHD) + disagg (on a website)

Suggestions for future studies

- Compare aggregate versus disagg (both on an IHD)
- Compare 2 groups:
  1. Aggregate on an IHD
  2. Aggregate (on an IHD) + disagg (on a website)
- Compare fine-grained disag versus coarse-grained disag

Suggestions for future studies

- Compare aggregate versus disagg (both on an IHD)
- Compare 2 groups:
  1. Aggregate on an IHD
  2. Aggregate (on an IHD) + disagg (on a website)
- Compare fine-grained disag versus coarse-grained disag
- If you have data then please consider releasing it; or writing a paper; or collaborating with someone who will write a paper with you!

Users might become more interested in disag feedback if:

Users might become more interested in disag feedback if:

- Energy prices increase

Users might become more interested in disag feedback if:

- Energy prices increase
- Concern about climate change deepens

Users might become more interested in disag feedback if:

- Energy prices increase
- Concern about climate change deepens
- Disag accuracy increases or if designers communicate uncertain estimates

Users might become more interested in disag feedback if:

- Energy prices increase
- Concern about climate change deepens
- Disag accuracy increases or if designers communicate uncertain estimates
- Lots of ideas in the literature about how to improve disag feedback. e.g. disag by *behaviour*; or display feedback near appliances; or provide better recommendations etc.

Appendix

PG&E 2014 trial results

Reported actions taken in response to feedback

Please describe the changes you or others in your household have made	Device Type		Stat. Sig. Diff (90%)?	P-value*
	Gateway	IHD		
Turned off lights not in use	75%	82%	Y	0.03
Turned off office equipment	44%	40%	N	0.26
Turned off entertainment center	32%	31%	N	0.87
Installed a power strip to control "vampire" loads	25%	18%	Y	0.02
Installed compact flourescent lights (CFLs)	20%	23%	N	0.48
Installed light-emitting diode lights (LEDs)	40%	38%	N	0.62
Bought an energy efficient appliance	15%	18%	N	0.24
Changed the setting on my manual thermostat to use less energy	16%	17%	N	0.57
Re-programmed by programmable thermostat to use less energy	22%	21%	N	0.63
Did fewer loads of laundry	32%	37%	N	0.10
Did fewer loads of dishes	20%	28%	Y	0.01
Only used cold water when doing laundry/dishes	15%	20%	Y	0.08
Other	21%	28%	Y	0.04

*p-values less than 0.1 indicate that gateway and IHD users' responses are significantly different at the 90% level of confidence.